

Forthcoming in *Biology & Philosophy*
Preprint (not copyedited or formatted)
DOI :10.1007/s10539-019-9689-8
Please use DOI when citing or quoting

Teleosemantics, Selection and Novel Contents

Justin Garson and David Papineau

Abstract: Mainstream teleosemantics is the view that mental representation should be understood in terms of biological functions, which, in turn, should be understood in terms of selection processes. One of the traditional criticisms of teleosemantics is the problem of *novel contents*: how can teleosemantics explain our ability to represent properties that are evolutionarily novel? In response, some have argued that by generalizing the notion of a selection process to include phenomena such as operant conditioning, and the neural selection that underlies it, we can resolve this problem. Here, we do four things: we develop this suggestion in a rigorous way through a simple example, we draw on recent neurobiological research to support its empirical plausibility, we defend the move from a host of objections in the literature, and we sketch how the picture can be extended to help us think about more complex “conceptual” representations and not just perceptual ones.

Introduction

Teleosemantics joins two ideas. First, it explains mental representation in terms of biological functions. Second, it explains biological functions in terms of natural selection, by equating functions with those effects for which biological traits are naturally selected.¹ Teleosemantics emerged in the 1980s as a leading naturalistic theory of intentionality and has been developed systematically since that time (see Papineau 2017; Neander 2017 for recent introductions).

Three main considerations support teleosemantics. First, it provides a naturalistic account of representation. To the extent that biological functions are part of the natural order, so too are representations. Second, it accounts for the possibility of misrepresentation, or representational error. A representational error, in this view, involves a failure of a trait to perform its biological function. Other naturalistic approaches to intentionality, such as informational (Dretske 1981), asymmetrical-dependence (Fodor 1990), or isomorphism theories (Cummins 1996) have struggled to account for misrepresentation. Third, teleosemantics is taxonomically general. It purports to account for representational capacities across the living world, and not just in humans and other cognitively sophisticated creatures. For example, it can help us understand what it is for a certain type of neural activation in an owl’s visual system to represent the location of a vole, or what it is for a vervet monkey to represent something as an approaching eagle.

¹ A few theorists explain representation in terms of biological functions, but adopt alternatives to the selected effects account of functions (examples include Abrams 2005 and Nanay 2013). We will understand ‘teleosemantics’ to exclude this option in what follows. We think that such views have trouble accounting for the possibility of malfunction (see next section).

Teleosemantics has its detractors. Some argue that natural selection is the wrong sort of thing to ground representation, reference, and ultimately, truth (e.g., Crane 2003; Burge 2010). After all, false beliefs and other misrepresentations sometimes promote reproductive success, just as true ones do. Others, such as Fodor (1990) and Strawson (2008), argue that teleosemantics faces intractable content indeterminacy problems. Natural selection cannot make the sorts of fine-grained property discriminations that a theory of content must deliver. We do not have space, here, to provide a comprehensive response to these challenges (but see Papineau 2017 and Neander 2017 for recent defences). Rather, we will focus on one particular problem, the problem of evolutionarily novel contents.

Here is the problem: at best, one might think, teleosemantics can help us understand how creatures are able to represent evolutionarily recurring features of their environments: how owls represent voles; how vervet monkeys represent approaching eagles; how humans represent dominance hierarchies, food sources, or mating opportunities. But it does not account for our ability to represent evolutionarily novel features of our environments, such as microwave heaters, traffic lights, or television celebrities. That is because, on the teleosemantic view, a representation's content is ultimately determined by the biological functions of the processes that involve it, which functions, in turn, depend on natural selection. So, teleosemantics seems overly limited in its ability to account the full spectrum of our representational capacities.

One initially plausible way of avoiding this problem is to recognize that biological natural selection, in sense of the intergenerational selection of genes², is not the only selection process in the natural world that creates new functions. The relevant notion of selection can be broadened to include other processes as well. For example, Millikan (1984), Papineau (1984, 1987, 1993), and Garson (2012) have argued that the trial-and-error shaping of behaviours and underlying neural structures should count as a function-bestowing selection process, because of its resemblance to natural selection in the evolutionary sense. These kinds of ontogenetic selection processes suggest one way in which teleosemantics can solve the problem of novel contents. If new *functions* can arise over the lifetime of the individual, and not just over an evolutionary time frame, then it seems plausible that new *representations* might arise over the lifetime of the individual as well.

In this paper, we defend the thesis that operant conditioning, and the neural selection processes that support it, can give rise to novel representations. Although the bare possibility of such a move has been stated in the literature, we develop it here in four specific ways. First, we develop the solution in a precise way using a simplified but true-to-life example of neural selection. We also show how the example “scales up” to let us

² There seems no principled reason for not including reliably vertically inherited non-genetic traits alongside genes as subject to intergenerational biological natural selection (e.g., Jablonka and Lamb 2005). We shall bypass this issue, however, given that it is orthogonal to our concerns.

attribute representational properties to more complex neural structures. Second, we rely on recent neurobiological research, specifically regarding the role of dopamine in instrumental conditioning, to support the biological plausibility of our thesis. Third, we defend the move from a host of objections, some from the existing literature and others novel. Some argue that the underlying theory of functions we accept is too liberal, some hold that single synapses cannot have representational properties, and some might maintain that the theory is redundant and therefore unnecessary. We think our theory survives these unscathed. Fourth, although we restrict our attention to perceptual representations, we sketch how the theory might be developed to help us think about conceptual representations as well. Here we rely heavily on Ruth Millikan's (2017) recent work on concepts. Though we have little systematic to say in this area, we believe our work shows the fruitfulness of our generalized selectionist approach. A significant virtue of our approach is that it can solve the problem of novel contents without introducing new, non-selectionist principles. Its frugality is a mark in its favour.

Here is how our article breaks down. We begin by developing the underlying view of function we accept to show that differential reproduction is not *necessary* for functions, but only differential persistence ("What are functions?"). We then show how our theory allows synapse selection to give rise to new representations, and how it "scales up" to cover more complex cases ("Synapse selection and novel representations"). We then turn to recent neurobiological research to support the plausibility of the basic picture we develop ("Is our thesis biologically plausible?"). We go on to defend the theory from a host of objections ("Objections and replies"). In the concluding section, we consider how we might extend the basic picture to other sorts of representations.

What are functions?

What are functions? In our view, the best theory of function is the selected effects theory. The selected effects theory holds that the function of a trait is, roughly, whatever it was selected for. For example, the function of the stag beetle's massive mandibles is to engage in male-to-male conflict, since that is what they were selected for by intergenerational selection. This is not the place to mount a full defence of the selected effects theory (see Garson 2016, Chapter 3, for an overview of and response to traditional criticisms), but it will be helpful briefly to rehearse its two main virtues: first, that it accounts for the explanatory aspect of functions, and, second, that it marks a clear distinction between function and malfunction.

First, functions are explanatory. One peculiar feature of functions is that, when biologists attribute a function to a trait, they are often trying to give a causal explanation for why those traits exist. One virtue of the selected effects theory is that it makes sense of this explanatory aspect of functions. A second peculiar feature of functions is that they carry with them a specification of malfunction. A trait malfunctions when it does not perform its function. This requires that a trait can *have* a function that it is nevertheless unable to *perform*. At first sight this can seem puzzling. How can something that is not done nevertheless be part of a trait's function? The selected effects theory, however, can easily accommodate this. For it holds that which function a trait *has* depends entirely on its

history. By contrast, whether or not a trait can *perform* that function depends entirely on its current structure and dynamics. So, there is no problem understanding how a trait can have a function it cannot perform. Tardive dyskinesia involves a dysfunction of the muscular system because the musculature is not able to do what it evolved by natural selection to do. We do not wish to say that no other theory of function can make sense of this, but such theories generally make heavy weather of the issue, to say the least (see Kingma 2010).

As noted above, some theorists have noted that the notion of selection, as understood in the selected effects theory, should not be understood narrowly to encompass only evolutionary natural selection, but that it should be expanded to include processes such as operant conditioning. Millikan (1984, p. 28), for example, suggests that behaviours can undergo a kind of differential reproduction that explains how they acquire new functions. We indicated our view that, if new *functions* can arise over the lifetime of the individual, then perhaps new *representations* could arise over the lifetime of the individual, too.

Kingsbury (2008), however, raises a challenge to the idea of functions that do not derive from intergenerational evolutionary selection. Proponents of the selected effects theory have often maintained that functions require something like differential reproduction or “copying.” Millikan (1984, p. 18), for example, says that functional traits must be members of “reproductively-established families.” Neander and Rosenberg (2012, p. 618) think that functions can only belong to what they call a “lineage of traits parsed by changes in the selection pressures operating on it.” If we accept this restriction, then any attribution of function will require the identification of something like replicators that form a lineage. As Kingsbury (2008, p. 502) has put it, “anyone who claims some process is a selection process needs to specify what it is, in that process, that reproduces and forms a lineage, and even given a very inclusive characterization of reproduction, this may not be easily done.”

To illustrate Kingsbury’s point, suppose one is entertaining two different hypotheses, H1 and H2, and then decides, after some reflection, to accept H1 and discard H2. H1 isn’t involved in anything like reproduction; it is simply *persisting* more or less successfully (p. 501). As Kingsbury says, “...to claim that a process is a selection process is to claim more than simply that it is a sorting process: it must also be a process which includes reproduction and differential reproductive success” (p. 505).

This objection arises in particularly acute form for our suggestion that novel representations come from ontogenetic neural selection. It is by no means obvious that anything worth calling *reproduction* occurs during neural development. While neural structures are differentially retained during development, there is no clear sense in which they replicate themselves. So our appeal to novel neural functions might seem to be in danger of falling at the first hurdle.

Our response to this challenge is to deny that reproduction (“copying”) is necessary for functions. We think that a trait can acquire a function simply through something like differential persistence in a population. In our view, restricting functions to entities that

reproduce is highly arbitrary. There seems no immediate theoretical motivation for this. Why not attribute functions to traits simply on the basis of differential persistence, that is, the ability of some traits to persist better (longer, more effectively) than others within a population?

In short, and pending some important qualifications, we suggest that the function of a trait is that effect that contributed, in the past, to the differential reproduction, *or* the differential persistence, of that trait (see Garson 2017 for more detailed discussion). The first part, “differential reproduction,” includes the sorts of entities that the traditional selected effects theory includes (such as the intergenerational selection of genes and perhaps antibody selection). The second part, “differential persistence,” includes processes such as neural selection. We will say a few more words by way of defence in the section entitled “Objections and replies.”

Let us make one final point about the contrast between differential persistence and differential reproduction. One might suspect that extending functions to include differential persistence is unnecessary. Perhaps one could think about neural selection as involving the differential *reproduction* of synapses. (See, for example, Artiga 2010, who suggests that various forms of ontogenetic learning fit the model of differential reproduction.) Consider the following scenario: suppose a neuron forms a synapse with another neuron. Suppose, as a result of that synapse, the first neuron is able to derive the nutritive support it needs to form new synapses. Shouldn’t we say that the later synapses are “copies” of the earlier synapse, or better, that there is a kind of parent-offspring relationship between the earlier and later synapses?

We think it’s a stretch to describe this case as involving reproduction or “copying” of synapses, for two reasons. In the paradigm cases of reproduction, there is a fairly tight causal connection between parent and offspring. Consider, for example, a cell that forms two daughter cells in mitosis. But the relationship between earlier and later synapses is highly indirect. An analogy can help. Suppose someone has a child, and the child brings her so much joy that she decides to have another. The first child is, in this indirect way, a cause of the second. But the second is not a reproduction or “copy” of the first; the parent-offspring relationship is entirely out of place here. We think the same point attaches to the differential proliferation of synapses.

An anonymous reviewer suggested a further way to challenge the idea that the later synapse is a “reproduction” of the earlier one: namely, there does not seem to be anything like *inheritance* of properties. Inheritance, here, would require the “winning” synapse to create synapses that are *more similar to it* than they are to “losing” synapse. We find no reason to think this would be the case, and so agree that this is another reason for not counting the proliferation of synapses as a case of differential reproduction, but as differential persistence instead.

Synapse Selection and Novel Representations

We have shown how functions merely require differential persistence, not reproduction.

The benefit of this insight is that we can account not only for the emergence of new functions, but also the emergence of new representations, over the individual's lifetime. This section will show how neural selection, by creating new functions, also creates new representations in the brain.

There are actually different types of neural selection, depending on the neural structure that undergoes competitive interactions: synapse selection, whole neuron selection, and, possibly neural group selection (Garson 2012). Here we will focus on synapse selection, both because it is empirically well-documented and because it is most directly involved in creating new representations.

The simplest possible case of synapse selection involves at least two synapses. Suppose there are two neurons, N_1 and N_2 , each of which synapse onto the same target neuron, N_T (see Figure 1). This yields two synapses, S_1 and S_2 . Suppose that, when S_1 is activated, it yields some useful downstream effect E . Synapse selection happens when, as a consequence of causing E , S_1 is retained and S_2 is eliminated.

In our view, if one synapse is retained over another because of some useful effect, then that synapse comes to have the *function* of producing that effect. Applied to our simple example, S_1 comes to have a novel function, namely, the function of producing E , for that is the activity that explains its differential retention in a population. (The *population* here consists of two synapses, S_1 and S_2 .) The case just described represents a “convergent” pattern of synapses, where two neurons innervate the same target. We could have just as easily illustrated the case through a “divergent” pattern of synapses, where a single neuron innervates two different targets.

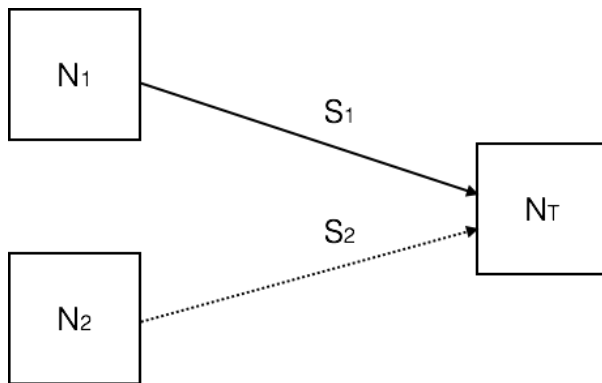


Figure 1. A simple case of synapse selection.

One core function of synapse selection is to help the individual adapt to the contingencies of its environment (Innocenti and Price 2005, p. 958). Consider how that might work in practice. (We have chosen a very simplified model for the purpose of illustration; soon we will show how the model “scales up” to accommodate more biologically realistic scenarios.) Suppose we place a rat into a new environment, one it has never been in before. There are two sorts of objects in that environment, red things and blue things. The red things are edible and the blue things are not. Suppose there is a certain neuron, N_1 ,

that tends to be activated by red things. Suppose there is another, N_2 , that tends to be activated by blue things. Suppose N_1 and N_2 both synapse onto the same target neuron, N_T , yielding S_1 and S_2 as in Figure 1. Whenever N_T is activated, it causes feeding behaviour in the rat. Now, suppose, on some occasion, the rat approaches a red thing; S_1 is activated, and the rat is rewarded with food. Suppose the rat approaches a blue thing; S_2 is activated; the rat tries to eat it, but there is no reward. Suppose, as a result, that S_1 is retained and S_2 is eliminated. Then, we wish to say that the *function* of S_1 is to cause feeding behaviour. For that is what it did that led to its differential retention in the relevant population (that is, its retention over other entities that were part of the same population, namely, the population consisting of S_1 and S_2).

We have seen how synapse selection creates new functions. But how does it create new representations? To see this, we should begin with the observation that a function of a trait is always some effect or another that it produces. However, sometimes a trait has the function of producing some effect *in the presence of* a certain stimulus or situation (Neander 2017 calls these “response functions”). For example, the pineal gland has the function of producing melatonin in response to the dimming of light.

When a state has the function of producing some behaviour when some condition obtains, teleosemantics allow us to say what content that state it has. The *content* of the state depends on what, in the world, it must correspond to in order for the resulting behaviour to be successful. The representational vehicle, here, is simply the activation of S_1 . S_1 has the function of producing feeding behaviour. So, what must S_1 correspond to in order for feeding behaviour to be successful?

The orthodox teleosemantic answer to this question is the presence of *something edible*. The function of the feeding behaviour is that the rat should ingest something nutritious. This will be achieved if the activation of S_1 is accompanied by the presence of something edible. Therefore (and setting aside various “content indeterminacy” worries), the activation of S_1 represents edible things (e.g., Millikan 1989b, p. 290).

Neander (2006, 2013, 2017) has developed a slightly different way of thinking about teleosemantics. In her view, the content of the representation does not depend as much on how the representation is utilized (“consumed”) by the system, but on how it is produced. So in our example she would say that S_1 represents, not edible things, but red ones. As she sees it, if the rat were to respond to a red thing that is not edible, S_1 would not be malfunctioning, but fulfilling its function of detecting red things. Neander agrees that the only reason that S_1 has the function of becoming activated in the presence of red things is that this disposition has been selected for enabling the rat to eat. Her view and the orthodox consumer view agree in their basic understanding of what the neural structure is selected for. In specific connection with attributing content, however, she places the emphasis on the producer mechanism rather than the consumer mechanism, partly

because she thinks that these content ascriptions are more relevant and helpful in the context of neuroscience.³

This is not the place to adjudicate between different versions of teleosemantics, specifically, the traditional consumer approach and Neander's producer approach. The more fundamental point is that the basic picture outlined here is consistent with the general way teleosemanticists ascribe content. In other words, whichever specific version of teleosemantics one chooses to adopt, all can agree that neural selection can give rise to new representations. Note in particular that nothing in these theories is inconsistent with some functions deriving from differential persistence rather than reproduction.

In our example we have focused on neural selection acting over single synapses. But nothing limits our basic picture of novel representations to single synapses. First, the basic model described here would apply just as well if S_1 were taken to represent a longer neural pathway involving many different synapses in succession. Second, the basic model would also work if, instead of only two neurons synapsing onto N_T , there were several neurons synapsing onto N_T , and some of them, as a group, are preferentially reinforced over others. Suppose a stimulus (such as a dog) causes the simultaneous activation of a group of synapses. Suppose the group of synapses that responds most vigorously to the stimulus is differentially strengthened, that is, strengthened over some other group that does not respond quite as vigorously (see Figure 1 of Wiggs and Martin 1998, p. 230, for discussion of how the selective elimination of groups of synapses might "sharpen" perceptual representations; also Wagner et al. 2004). In that case, we may treat the former group of synapses as a unit for the purpose of assigning a function to it. Nothing in our basic account limits us to attributing functions, and hence contents, to single synapses.

We will make two more points of clarification before moving on. The first has to do with the nature of the representational vehicle involved here. What is it, precisely, in the brain that carries semantic content? In the simplest possible scenario, where S_1 has been selected over S_2 , the vehicle is S_1 's *being activated*: specifically, the release of neurotransmitter into the synapse (or other signalling molecules as in gap junctions). However, in more complex cases, the representational vehicle could be something more sophisticated. For example, suppose there are two neurons, N_1 and N_2 , both of which synapse onto the same target N_T and both of which are responsive to the same external

³ We have reservations about Neander's way of carrying out the teleosemantic project. We think she has problems with *distal content*. For Neander, some neural activity R (for example, some optic nerve activity in the frog) represents some external stimulus S (a fly) so long as there is a mechanism that has the evolved function of producing R in response to S . But consider some inner, proximal stimulus, S^* , that forms a midway point in the causal chain between S and R , for example, the retinal impression created by the fly. Then, presumably, there is an evolved mechanism that has the function of producing R in response to S^* , too. But intuitively, R is about S , and not *also* about S^* . Neander's proposed solution (2013, 2017) in terms of asymmetric "informational" relationships between various stimuli has recently come under fire from Schulte (2018). One benefit of relying on consumer-based teleosemantics is that it skirts this problem entirely.

stimulus. Suppose that N_1 responds with a much more rapid firing rate (measured by spike count over a certain window of time) than N_2 , and as a consequence, S_1 is retained over S_2 . Then we should say that the representational vehicle, in this case, is a spike train with spike count n .

We have explained how behaviour-relevant neural synapses can come to represent novel environmental conditions as a result of ontogenetic neural selection. This shows that there is at least one way in which teleosemantics can account for novel representational content. However – and this is our second point of clarification – we certainly do not want to imply that all cases of neural representation of environmental conditions must involve such ontogeny-based selection. It is perfectly possible that many similar behaviour-relevant neural synapses are “hard-wired” because of intergenerational genetic selection, in which case they will represent in an entirely analogous way.

For example, consider that such genetic selection has come to ensure that certain groups of optic fibres in the frog are activated by bug-like stimuli, and that such stimuli prompt tongue-snapping behaviour (see Lettvin et al. 1957, who called these optic fibres the “bug-perceivers”). Those hard-wired pathways will have the selection-based function of prompting that behaviour in those circumstances, and so, according to teleosemantics, will represent the presence of an insect (or possibly something with an insect-like appearance). In this kind of case, however, the function will be based on intergenerational genetic selection, rather than on ontogenetic neural selection, and the represented condition will be one present in the ancestral environment, and not just in the organism’s individual experience.

Interestingly, there is reason to suppose that the genetic selection of hard-wired neural connections will often be a consequence of prior ontogenetic development of just those same connections. Genetic selection on its own can have trouble evolving structures whose functionality requires a number of different genetically controlled components to be in place. The co-occurrence of the relevant genes will be rare, and in any case will be broken up in sexual species by genetic recombination. But if such structures can initially be developed by ontogenetic learning, then any one of the relevant genes will have a selective advantage on its own, since it will make it easier for the organism to learn the other components. This suggests that the genetic shaping of complex neural structures might be driven by a kind of progressive “Baldwinization,” in which proto-versions of the structures are first learned, and then come increasingly under genetic control, which then allows more complex versions to be learned, which in turn also come under genetic control, and so on (Papineau 2005, 2006).

Is our thesis biologically plausible?

So far, we have simply illustrated how an instance of synapse selection might give rise to novel representations. But does the brain really work that way? Recent research into the neural foundations of operant conditioning supports the basic picture presented here, by suggesting that dopamine acts as a reinforcement signal that selectively strengthens certain synapses *because* of their rewarding consequences. We shall briefly outline the

main features of this research.

Brain scientists have long suspected that dopamine plays a role in operant conditioning. In the 1950s, work by the neuroscientists Olds and Milner (1954) showed that electrical stimulation of certain brain regions could cause rats to vigorously self-stimulate. At the time, it was unknown which cortical areas and neurotransmitters were involved. More recent research using ontogenetic methods has clarified the issue, by showing that it is specifically the stimulation of dopamine neurons in the ventral tegmental area (VTA) of the midbrain that causes rats to self-stimulate (Witten et al. 2011).

As to the precise function played by dopamine in the VTA, one widely accepted hypothesis is that dopamine signals not rewards per se, but *unexpected* rewards (and punishments). This is the “reward-prediction error” hypothesis, which holds, in part, that an unexpected reward causes a phasic firing of VTA dopamine neurons which release dopamine onto limbic structures, specifically, the nucleus accumbens, as well as neurons in the prefrontal cortex. (See Schultz et al. 1997; Schultz and Dickinson 2000; also see Columbo 2014 for a recent philosophical discussion.)

The connection between this phasic VTA firing and reinforcement of behaviours is less clearly understood, but one plausible idea is that the phasic release of dopamine acts as a kind of global signal that *selectively reinforces just those synapses that were involved in the rewarding behaviour*. As Schulz and Dickinson (2000, p. 490) put it, “with competing synaptic inputs, neuronal activity occurring simultaneously with the error signal would be preferentially processed. This mechanism would result in the biasing, prioritization, or selection of certain inputs over others...” Schultz et al. (1997, p. 1598) explicitly describe the reinforcement of certain synapses over others as a sort of “competition:” “the dopamine signal influences behavioural choices by modulating the level of competition in the dorsal striatum.”

Which synapses, precisely, are undergoing selection? The relevant synapses connect structures in the prefrontal cortex to the nucleus accumbens (corticostriatal synapses). Possibly, these synapses carry information about the sensory features of the stimulus to the nucleus accumbens, where they can affect behaviour. In our view, if one corticostriatal synapse is “preferentially processed” over another, *because* it helps generate appropriate behaviour in the presence of the stimulus, then its *function* is to help generate that behaviour in the presence of the stimulus. In short, it is plausible that synapse selection is actually involved in operant conditioning. (For more discussion on structural changes in corticostriatal neurons in learning and addiction, also see Robinson and Kolb 1997; Hyman and Malenka 2001; Kauer and Malenka 2007.)

Objections and replies

In this section we will consider four main objections to the picture of novel contents developed above. The first objection is due to Kingsbury (2008), who argues that *differential persistence* is far too liberal a basis for novel functions. The second is due to Cao (2012), who maintains that we should not think of the activation of single neurons,

or single synapses, as carrying information, but only large collections of neurons. Third, we consider whether our theory has the counterintuitive consequence that associative learning cannot create new representations. Finally, we consider whether our appeal to ontogenetic selection processes is redundant, given the prospect that Millikan's distinction between *direct* and *derived* proper functions can do the needed work.

Objection 1: Is differential persistence too liberal for functions?

Kingsbury considers, then rejects, one of the crucial moves we make here. We argue that functions merely require a population of entities that undergo something like differential persistence; differential reproduction is not necessary. She argues that such a move would render the underlying theory of function overly liberal to the point of vacuity. She asks us to imagine a bunch of rocks on a beach that vary in their hardness. Some rocks, because of their hardness, erode less quickly than others. So the rocks persist differentially because of their hardness. But surely we do not want to say that hardness is a function of rocks (Kingsbury 2008, p. 496; also see Lewens 2004, p. 127 for similar thoughts).

We agree that hardness is not a function of rocks. We also agree that one way to avoid that implication is by restricting functions to things that reproduce. But that is not the only way to avoid that implication. We think that there are at least two other promising options, one that appeals to the nature of *populations*, and one that hinges on the requirements for *cumulative* selection.

First, natural selection always takes place in a population (e.g., Millstein 2009; Godfrey-Smith 2009), and it is widely agreed by philosophers of biology that a population must involve some form of fitness-relevant interactions between its members (Godfrey-Smith 2009, p. 52; Millstein 2009, p. 271; Matthewson 2015, p. 180). *My* behaviour must have some effect on *your* fitness, or vice versa. In our view, groups of neural structures can form populations, but collections of rocks cannot. That is because, unlike a collection of rocks, neural structures can be involved in competitive interactions that affect each others' fitness. (In line with this, we deny that "fitness-relevant interactions" requires *reproduction*; we read "fitness-relevant interactions" broadly enough to encompass interactions that affect a neural structure's prospects for survival.) The crucial point for our purposes is that a group of rocks on a beach clearly fails the requirement, since the rate at which one rock erodes has no effect on the rate at which others erode (Garson 2017).

Let us now consider the second way to block the attribution of functions to rocks. This invokes the requirements for cumulative selection. In the standard biological examples, natural selection has a cumulative character. Adaptations build on adaptations to form extremely complex and sophisticated traits, such as the human eye. But there is nothing like cumulative selection taking place in our pile of rocks. It is not as if those rocks that are "selected" on account of their hardness then go on to participate in new selection processes for other traits. As Kingsbury (2008, p. 6) puts it, the example merely involves a "one-shot sorting" process.

There are principled reasons for attaching importance to cumulative selection. Some philosophers of biology hold that natural selection never explains the origin of traits, but only their retention (the so-called “negative view”). In their view, origins always result from mutations, and the role of selection is merely to decide which mutations escape elimination (Sober 1984, 1995). Karen Neander has responded that this argument only applies to “one-shot” selection processes. When it comes to complex products of cumulative selection, like the human eye, then selection clearly plays a role in building traits, as well as preserving them. It is only after selection has first fixed the primitive stages of the trait in the population that it is possible for later selection to add on improvements (Neander 1995; see Godfrey-Smith 2009, pp. 48-53).

Neural selection can also involve a version of cumulative selection. Consider the well-studied retinal ganglion cells (RGC) of the zebrafish, which form synapses with the optic tectum. Recent evidence suggests that when two or more retinal axons are competing over tectal sites, the “winning” axon not only gains control over a single neuron, but it can send forth new branches (axonal arborization) and make new synaptic partners (Hua et al. 2005; Gosse et al. 2008; see Kalil and Dent 2014 for discussion). In theory, the “winning” neuron could go on to engage in further rounds of competitive interactions. Cumulativity is thus apparent at the neural level (see Czikó 1995, pp. 64-69 for more on the prospect of cumulative neural selection).

Perhaps there are reasons for *not* restricting functions to traits that undergo cumulative selection. Consider the famous peppered moth (*Biston betularia*) in Manchester in the nineteenth century. In the 1840s, the dark form represented a small fraction of the population relative to the white; by the 1890s, largely as a result of industrial pollution, it spread to more than 90% of the population. Here, the relevant selection involves a fairly simple trait, wing colour, and a fairly simple activity, camouflage. It’s not clear, at the outset, that there’s anything like cumulative selection going on. Nonetheless, many would say the dark wing has the function of camouflage.⁴

One possible move is to distinguish between sorts of functions that do and do not require cumulative selection. For example, Allen and Bekoff (1995, p. 615) introduce the notion of *natural design* to describe the cumulative cases where natural selection actually changes the structure of some trait, rather than simply changing trait frequencies, while allowing that the simpler non-cumulative process can give rise to functions.

If we do adopt this more permissive line on functions, then we will not be able to rule hardness of the rocks non-functional on the grounds that it lacks the right kind of cumulative provenance. Still, there remains the point that the rocks do not form an

⁴ Similarly, not all neural selection need be cumulative. Consider synapse selection in the neuromuscular junction in rats. At birth, several motor neurons synapse onto each muscle fiber; within several weeks, only one remains. Here, the “winning” neuron does *not* typically go on to make new synaptic partners, but we would still be happy to say that its synapse acquires a new function via differential persistence.

interacting population, which in any case gives us reason to deny that they are subject to a function-bestowing selection processes.

All in all, then, there is plenty of room to block Kingsbury's worry that trivial counterexamples will crowd in once we take the view that non-reproducing entities, like behaviours and neural structures, can acquire functions by selection of effects. The cases we are interested can be differentiated from the rocks in two ways that are relevant to the attribution of functions: they involve cumulative selection, and this selection operates on genuine populations of interacting individuals. The rocks' lack of cumulative selection certainly means that their hardness is not "designed," in the way that complex traits are. And while this might arguably leave open that the rocks have some minimal level of functionality, along with the peppered moths, even this much is ruled out by the lack of an interacting population.

Objection 2: Can single neurons or synapses have representational properties?

We stated that if synapse S_1 is selected over synapse S_2 by virtue of producing some useful behaviour in response to a certain stimulus, then the activation of S_1 constitutes a representation, and its content is whatever in the world S_1 's consumers need to discharge their functions in the historically normal way. For example, in our example of the rat with edible foods, the orthodox teleosemantic interpretation of S_1 's activation is that *there is something edible*. Cao (2012), however, has argued that we should generally avoid attributing contents to single neuron activations (such as a burst of neurotransmitter in a synaptic cleft), but only to vast ensembles of neurons or even to the whole organism. The reason, she thinks, is that the informational content that one neuron can transmit to another is too sparse or uninformative. As she puts the point, "small groups of cells only have minimal informational access to anything beyond the activities of other groups of cells nearby" (p. 70). For example, a dendrite binding to neurotransmitter only has access to information such as, "[There is] X quantity of neurotransmitter in my cleft" (p. 65).

We disagree with her assessment of content. In particular, we think that Cao assumes, without sufficient evidence, that the content of a single neuron's activation must be something highly *proximal* to that neuron. To illustrate the point, suppose there are two neurons in succession: N_1 and N_T , as in Figure 1. Suppose an external stimulus (an edible berry) causes N_1 to fire, which causes N_T to fire, which ultimately causes feeding behaviour. What is the content of N_1 's firing, that is, activation of the N_1 - N_T synapse? She thinks it must be something like, *N_1 is presently active*. As Neander (2017, p. 137) would put it, Cao assumes that N_1 - N_T synaptic activation must represent its most proximal *normal cause*. But we do not see why, on her account, a burst of neurotransmitter can only be about its own most proximal normal cause.

Not only do we not see how that result follows from her account, but the standard consumer-based approach to teleosemantics actually implies something quite different, namely, that the release of neurotransmitter has the distal content, *there is something edible*. That is because, on the standard approach, the content of a representation is whatever it must correspond to for the consumer mechanisms to perform their function in

the historically normal way. Assuming that the ultimate consumer mechanism is, here, a motor routine designed to elicit feeding behaviour, then, in order for that mechanism to discharge its function, there must in fact be something edible in its environment.

Objection 3: Does our view let associative conditioning create new representations?

Not all cases of novel representational contents resulting from neural development will fit our model. In particular, consider classical associative conditioning, or Hebbian learning. This involves the reinforcement of synaptic connections between neurons that have previously fired together (Kandel et al. 2013, p. 1467). A dog hears a bell repeatedly preceding the arrival of food, and the neurons prompted by the bell will become connected to those that register the presence of food. We certainly wish to allow this sort of learning to give rise to new representations, too.

Suppose we have some organism that is already capable of representing some A and B (bells and food say) as a result of either ontogenetic or phylogenetic selection. If it is then classically conditioned to anticipate B after detecting A, we can view the synaptic connection then formed as representing that *A will be followed by B*. This will be a novel representational content for the organism, for there will be no reason to suppose that As were followed by Bs in its ancestral environments. Note that in this case, however, the new synaptic representations will not depend on ontogenetic *selection* for some effect. There is no feedback loop that reinforces the synapses just in case some effect occurs. Rather the mechanism is simply that neurons that fire together will come to be wired together, quite independently of any effects.

This is neither to say that the classically conditioned synapses do not serve functions, nor that their novel representational contents cannot be analysed teleosemantically in terms of those functions. These functions will not depend on ontogenetic selection, however, but will be *derived* functions, in Millikan's sense (Millikan 1984, p. 42; also see her 1989a, p. 288). Millikan's notion of a derived function hinges on the point that some genetically evolved mechanisms have the function of producing traits that will serve further functions in relation to their current environments. The chameleon's camouflage mechanism has the function of producing skin colouring that will hide the chameleon against the current background. As a result, the current skin colour will have the "derived proper function" of hiding the chameleon against, say, the current yellow background. But this specific derived function need not be one that was ever served in the phylogenetic past. That particular shade of yellow might never have occurred in ancestral chameleon environments.

Similarly, we can suppose that the classical conditioning mechanism has the function of enabling organisms to form neural associations that will represent whatever regularities happen to be displayed in their current environments. This now gives us another way in which novel contents can be accommodated by teleosemantics. As well as novel contents resulting from ontogenetic selection, as outlined in this paper so far, there are also cases where neural structures represent novel environmental features in virtue of their derived functions (Kingsbury 2006 develops this theme). In these latter cases, the selection

processes giving rise to the relevant functions are all phylogenetic, rather than ontogenetic. Intergenerational genetic selection has constructed a learning mechanism that, without the help of any further ontogenetic selection, produces traits with current-environment-related functions.

Objection 4: Is our appeal to ontogenetic selection processes redundant?

The previous subsection raises an obvious question about our emphasis on ontogenetic selection in the main body of this paper. Do we need to invoke neural selection to account for the functions resulting from operant conditioning? After all, there is the option of viewing the neural structures resulting from operant conditioning as having derived functions in Millikan's sense, just as we did with the products of classical conditioning.

That is, we could attribute their functionality to their deriving from a mechanism – operant conditioning – that itself results from intergenerational genetic selection. The function of this operant conditioning mechanism is to produce neural structures whose function, in turn, is to adapt the organism to the contingencies of its local environment. So viewed, the functions that these novel neural structures acquire during the organism's lifetime (e.g., to produce feeding behaviour in the presence of red berries), would derive from the intergenerational selection of genes, independently of the fact that ontogenetic neural selection happens to be involved in their development.

A first point to make here is that, even if we accepted this analysis, it would not affect our main point. We have been concerned to show how teleosemantics can explain the representation of circumstances that are evolutionarily novel. To do so we have observed that neural selection results in structures with the function of gearing behaviour to evolutionarily novel circumstances, which structures therefore represent those circumstances. This point will stand even if we see such functionality as deriving ultimately from phylogenetic selection, rather than proximally from neural selection.

Having said that, we think it is problematic to explain the functionality of these ontogenetically acquired neural structures in terms of derived proper functions. As we noted in the second section, "What are functions?," the traditional selected effects theory unnecessarily, and we think arbitrarily, limits ("direct proper") functions to traits that undergo something like differential reproduction or "copying." In the present context, this arbitrary restriction is what motivates the idea that phylogenetic natural selection alone creates direct proper functions, and that neural selection, at best, creates derived proper functions. We see no reason to privilege phylogenetic natural selection in this way. If the ancestral selection of differentially reproducing genes can underpin functions, why deny the same functional significance to the ontogenetic selection of differentially persisting synapses?

Perhaps this is not an issue worth fighting over. As it happens, all familiar examples of ontogenetic selection mechanisms, such as operant conditioning and the adaptive immune system (e.g., Hull et al. 2001) are also mechanisms that have been designed by

phylogenetic selection to produce derived functional traits. So, as we observed above in connection with neural representation, the traits will come out as having the same functions whether or not we view ontogenetic selection as creating functions in its own right.

But, just to pursue the point, perhaps there are some few special cases which can focus the issue and show that differential persistence selection mechanisms can create functions even if they are not the upshot of prior phylogenetic biological design. For example, consider the way in which operant conditioning mechanisms can be “hijacked” by the ingestion of addictive substances.

It is widely accepted that phasic dopamine functions as a reward prediction error signal in operant conditioning mechanisms (see “Is our thesis biologically plausible?”). When behaviour leads to biologically rewarding outcomes, then the reward triggers the release of dopamine and leads to the reinforcement of relevant synapses. However, the ingestion of certain substances, including amphetamines, alcohol, nicotine, opiates, cocaine, and benzodiazepines, causes high levels of dopamine release independently of leading to reward, and so “short-circuits” the normal reinforcement mechanism to encourage behaviours that lead to ingestion of such dopaminergic substances (Butlin and Papineau 2015). (Opiates play a slightly more complicated role in addiction since, in addition to their ability to trigger dopamine release, they can also directly target opioid receptors in nucleus accumbens neurons – see Hyman and Malenka 2001, p. 679.)

We can view this short-circuited system as a selection mechanism in its own right, one which favours the persistence of just those synaptic connections that lead to dopamine release. And, in line with this, it seems natural to say that the behaviours of subjects who have been conditioned by this system, as well as the neural structures that facilitate those behaviours, have the *function* of leading them to seek alcohol, or cocaine, or whichever dopaminergic substances were the means to such release.

But note that these functions are not derived functions resulting from the operation of the normal operant conditioning system. The latter system has the function of generating behaviours that in the individual’s experience have led to *biologically rewarding* outcomes. This is not true of behaviours resulting from the dopamine short-circuit. They are being generated by a “rogue” selection mechanism, an unintended by-product of the normal operant conditioning system. In the presence of dopaminergic substances⁵, this

⁵ In a sense, the “rogue” system will also be present and operating even in the normal cases where biologically rewarding outcomes precede dopamine release. Note that, where the biologically designed conditioning system selects *for* rewarding outcomes, with dopamine release merely the procedure it uses to reinforce synapses, the rogue system, by contrast, selects *for* dopamine release itself. In most cases, the rogue system will be intertwined with the biological one, they will select the same behaviours, and the non-standard functionality deriving from the rogue system will go unremarked. It is specifically those cases where the rogue system operates without the biological one that highlight the non-standard functionality of its selected behaviours.

rogue system selects behaviours that do not lead to biologically rewarding outcomes, but directly to dopamine release. Their reinforcement properties occur independently of the subjective effect of the drug or other aspects of the experience (Hyman et al. 2006, p. 575).

Still, as we observed, it seems natural to say that such behaviours have *functions* that relate to their rogue selection, of leading the user to seek alcohol, or cocaine, or whatever, notwithstanding the fact that these functions are not in line with phylogenetic biological design. We take this to support the view that the ontogenetic selection of differentially persisting items can create functions independently, even in cases where these are not also derived from mechanisms shaped by phylogenetic selection.

Conceptual representations

We have illustrated the possibility of ontogenetically selected novel contents for very simple “non-conceptual representations” of conditions like *red edible thing now here*. Readers might wonder how our story might generalize to more complicated cases.

This has been an ongoing issue for teleosemanticists. One strategy would be to adopt a multi-tiered approach that appeals to selection only at an initial stage to explain ground-level perceptual representations, and then works its way up to conceptual representations by introducing fundamentally different, and non-selectionist, kinds of processes. (See e.g. Sterelny 1990, pp. 138-40, Neander 2017, pp. 205-15.)

We are not convinced, however, that one must move away from selectionist processes to understand more high-level, conceptual representations. In humans and other cognitively sophisticated animals, selection processes can be invoked to explain not only ground-level perceptual representations, but higher-level conceptual representations as well. Millikan’s (2017) recent work has done much to point the way. The ability to refer to and re-identify spatiotemporal particulars, biological taxa, chemical kinds and other “substances” allows cognizing subjects to lock onto rich sources of usable information. Our cognitive powers are designed accordingly. As we encounter new substances in our individual lifetimes, we readily coin concepts whose purpose is track and accumulate information about those substances.

A bit of background: Millikan urges us to dispense with the idea of *concepts*, which are supposed to be shared between people, and to adopt the language of “unicepts” and “unitrackers” instead. *Unicepts* are like concepts, in that they compose propositions, but they are unlike concepts in that they generally will not be shared between any two individuals – they are idiosyncratic. Each unicept comes paired with a corresponding *unitracker*. A unitracker is a mechanism that has the function of re-identifying individuals and kinds. My unitracker for the kind *dog* can be thought of as a mechanism that has the function of scanning the environment for instances of dogs, and tokening a judgment involving the corresponding unicept (*there is a dog*). The referent of a unicept is just the target of its corresponding unitracker (p. 44). My unicept for dog refers to dogs because it is paired with a unitracker that has the function of tracking dogs.

So how do unitrackers get their functions? What determines what a unitracker is “supposed to” track? Some unitrackers might have their targets innately, because they have been shaped by natural selection by virtue of their differential success in tracking some evolutionarily recurring feature of their environments. In some cases, where the referents are of types that have been encountered in ancestral environments, one can imagine “pre-designed” concepts which lock onto entities like individual humans or biological species by a mechanism akin to imprinting (p. 74). But that would still leave open the question of how there can be unitrackers for entirely novel substances and attributes: for traffic lights, or cotton candy, or the property of having a brawny aftertaste or three-day stubble. As Millikan puts it, “...how could a totally new kind of unitracker emerge, one that is not supplied in advance with any clear method of priming or any template” (p. 79)?

In these cases, she suggests that some element of trial-and-error learning will be involved, with unitrackers being retained or discarded depending on whether they are *nonempty* and *univocal*, that is, depending on whether they track one and only one entity. Crucially, the test for such unitrackers is that they generally yield judgments that conform to the law of noncontradiction. In short, unitrackers undergo a selection process in which some are retained and some are discarded on the basis of their conformity to said law: “This proxy selection mechanism uses regular conformity of unitracker outputs with the laws of identity and noncontradiction as its selection criterion. It discards, retains, or alters unitrackers according to their reliability in helping to confirm and reconfirm the same judgments...without also producing contrary judgments, judgments that contradict” (p. 80).

To use a simple example: suppose someone develops a unitracker that gets “triggered” by two kinds of things: cotton candy and fiberglass insulation. He has a corresponding unicept, *R*. (So, when he encounters cotton candy or fiberglass, he entertains something like the thought, *there is another R*, or *there is that R stuff again*.) At one time, he eats a piece of cotton candy and forms the judgment that the substance tracked by the unitracker has a sweet taste: *R is sweet*. Later, he eats a piece of fiberglass and forms the judgment that it is not sweet: *R is not sweet*. The unitracker has produced a contradiction, so he either eliminates it entirely, or investigates these substances more carefully as a prelude to refining his unitracker or splitting it into two. This would be a case in which representations *emerge*, or are *refined*, through an ontogenic selection process. Again, our point is not to endorse Millikan’s specific views about how novel concepts get off the ground. The point is to gesture toward how the selectionist mechanism we describe here might be extended to help us think about conceptual representations and not just perceptual ones.

We hope that this paper has made it clear how ontogenetic selection mechanisms are a legitimate resource for teleosemantics. In particular, we think our appeal to ontogenetic mechanisms represents a sharp break from how many teleosemanticists have tried to solve the problem of novel contents. As noted above, a standard view in the literature is a “two-tiered” view; selection is invoked to explain how evolutionarily archaic

representations acquire their contents, and then other, non-selectionist principles are invoked to explain how we represent novel environmental properties. Though we do not dismiss those attempts outright – they might indeed yield fruitful insights for thinking about content – we reject the need for this bifurcated model in explaining novel contents. We think the parsimoniousness of our view is one of its virtues.

References

- Abrams, M. 2005. Teleosemantics without natural selection. *Biology and Philosophy* 20: 97-116.
- Allen, C., and Bekoff, M. 1995. Biological function, adaptation, and natural design. *Philosophy of Science* 62: 609-622.
- Artiga, M. 2010. Learning and selection processes. *Theoria* 68: 197-209.
- Burge, T. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.
- Butlin, P., and Papineau, D. 2015. Normal and addictive desires. In *Addiction and Choice*, eds. N. Heather and G. Segal, 99-115. Oxford: Oxford University Press.
- Cao, R. 2012. A teleosemantic approach to information in the brain. *Biology and Philosophy* 27: 49-71.
- Columbo, M. 2014. Deep and beautiful: The reward prediction error hypothesis of dopamine. *Studies in History and Philosophy of Biological and Biomedical Sciences* 45: 57-67.
- Crane, T. 2003. *The Mechanical Mind (2nd ed)*. London: Routledge.
- Cummins, R. 1996. *Representations, Targets, and Attitudes*. Cambridge, MA: MIT Press.
- Cziko, G. 1995. *Without Miracles: Universal Selection Theory and the Second Darwinian Revolution*. Cambridge, MA: MIT Press.
- Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.
- Fodor, J. A. 1990. *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.
- Garson, J. 2012. Function, selection, and construction in the brain. *Synthese* 189: 451-481.
- Garson, J. 2016. *A Critical Overview of Biological Functions*. Dordrecht: Springer.
- Garson, J. 2017. A generalized selected effects theory of function. *Philosophy of Science* 84: 523-543.
- Godfrey-Smith, P. 2009. *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.
- Gosse, N. J., et al. 2008. Retinotopic order in the absence of axon competition. *Nature* 452: 892-895.

- Hua, J. Y., et al. 2005. Regulation of axon growth *in vivo* by activity-based competition.
- Hull, D. L., Langman, R. E., and Glenn, S. S. 2001. A general account of selection: Biology, immunology and behavior. *Behavioral and Brain Sciences* 24: 511–527.
- Hyman, S. E., et al. 2006. Neural mechanisms of addiction: The role of reward-related learning and memory. *Annual Review of Neuroscience* 29: 565-598.
- Hyman, S. E., and Malenka, R. C. 2001. Addiction and the brain: the neurobiology of compulsion and its persistence. *Nature Reviews Neuroscience* 2: 695-703.
- Innocenti, G. M., and Price, D. J. 2005. Exuberance in the development of cortical networks. *Nature Reviews Neuroscience* 6: 955-65.
- Jablonka, E., and Lamb, M. J. 2005. *Evolution in Four Dimensions*. Cambridge, MA: MIT Press.
- Kalil, K, and Dent, E. W. 2014. Branch management: Mechanisms of axon branching in the developing vertebrate CNS. *Nature Reviews Neuroscience* 15: 7-18.
- Kandel, E. R., et al. 2013. *Principles of Neural Science, 5th ed*. New York: McGraw Hill.
- Kauer, J. A., and Malenka, R. C. 2007. Synaptic plasticity and addiction. *Nature Reviews Neuroscience* 8: 844-858.
- Kingsbury, J. 2006. A proper understanding of Millikan. *Acta Analytica* 21: 23-40.
- Kingsbury, J. 2008. Learning and selection. *Biology and Philosophy* 23: 493–507.
- Kingma, E. 2010. Paracetamol, poison, and polio: Why Boorse’s account of function fails to distinguish health and disease. *British Journal for the Philosophy of Science* 61: 241–264.
- Lettvin, J. Y., et al. 1959. What the frog’s eye tells the frog’s brain. *Proceedings of the IRE* 47: 1940-1951.
- Lewens, T. 2004. *Organisms and Artifacts: Design in Nature and Elsewhere*. Cambridge, MA: MIT Press.
- Matthewson, J. 2015. Defining paradigm darwinian populations. *Philosophy of Science* 82: 178-197.
- Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

- Millikan, R. G. 1989a. In defense of proper functions. *Philosophy of Science* 56: 288-302.
- Millikan, R. G. 1989b. Biosemantics. *Journal of Philosophy* 86: 281-297.
- Millikan, R. G. 2017. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford: Oxford University Press.
- Millstein, R. L. 2009. Populations as individuals. *Biological Theory* 4: 267–273.
- Nanay, B. 2013. Success semantics: The sequel. *Philosophical Studies* 165: 151-165.
- Neander, K. 1995. Pruning the tree of life. *British Journal for the Philosophy of Science* 46: 59-80.
- Neander, K. 2006. Content for cognitive science. In *Teleosemantics*, ed. G. Macdonald and D. Papineau, 167-194. Oxford: Clarendon Press.
- Neander, K. 2013. Toward an informational teleosemantics. In *Millikan and her Critics*, ed. D. Ryder, J. Kingsbury, and K. Williford, 21-36. Malden, MA: Wiley-Blackwell.
- Neander, K. 2017. *A Mark of the Mental: In Defense of Informational Teleosemantics*. Cambridge, MA: MIT Press.
- Neander, K., and Rosenberg, A. 2012. Solving the circularity problem for functions. *Journal of Philosophy* 109: 613-22.
- Olds, J., and Milner, P. 1954. Positive reinforcement produced by electrical stimulation of the septal area and other regions of the rat's brain. *Journal of Comparative and Physiological Psychology* 47: 419-427.
- Papineau, D. 1984. Representation and explanation. *Philosophy of Science* 51: 550-72.
- Papineau, D. 1987. *Reality and Representation*. Oxford: Blackwell.
- Papineau, D. 1993. *Philosophical Naturalism*. Oxford: Blackwell.
- Papineau, D. 2005. Social learning and the Baldwin effect. In *Evolution, Rationality and Cognition*, ed. A. Zilhão, 40-60. London: Routledge.
- Papineau, D. 2006. The cultural origins of cognitive adaptations. In *Philosophy, Biology, and Life*. Ed. A. O'Hear, 291-324. Cambridge: Cambridge University Press.
- Papineau, D. 2017. Teleosemantics. In *How Biology Shapes Philosophy*, ed. D. L. Smith, 95-120. Cambridge: Cambridge University Press.

- Robinson, T. E., and Kolb, B. 1997. Persistent structural modifications in nucleus accumbens and prefrontal cortex neurons produced by previous experience with amphetamine. *Journal of Neuroscience* 17: 8491-8497.
- Schulte, P. 2018. Perceiving the world outside: How to solve the distality problem for informational teleosemantics. *The Philosophical Quarterly* 68: 349-369.
- Schultz, W., et al. 1997. A neural substrate of prediction and reward. *Science* 275: 1593-99.
- Schultz, W., and Dickinson, A. 2000. Neuronal coding of prediction errors. *Annual Review of Neuroscience* 23: 473-500.
- Sober, E. 1984. *The Nature of Selection*. Chicago: University of Chicago Press.
- Sober, E. 1995. Natural selection and distributive explanation: A reply to Neander. *British Journal for the Philosophy of Science* 46: 384-397.
- Sterelny, K. 1990. *The Representational Theory of Mind*. Oxford: Blackwell.
- Strawson, G. 2008. Real intentionality 3: Why intentionality entails consciousness. In *Real Materialism and Other Essays*, ed. G. Strawson, 281–305. Oxford: Oxford University Press.
- Wagner, A. D., Bunge, S. A., and Badre, D. 2004. Cognitive control, semantic memory, and priming: Contributions from prefrontal cortex. In *The Cognitive Neurosciences, 3rd ed*, Ed. M. Gazzaniga, 709–726. Cambridge, MA: MIT Press.
- Wiggs, C. L., and Martin, A. 1998. Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology* 8: 227-233.
- Witten, I. B., et al. 2011. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* 72: 721-33.